B3

**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
A61K 37/02, C07K 7/00, 13/00
C07K 15/00

A1

(11) International Publication Number: WO 92/07579

(43) International Publication Date: 14 May 1992 (14.05.92)

(21) International Application Number: PCT/US91/08320

(22) International Filing Date: 6 November 1991 (06.11.91)

(30) Priority data:
609,915 6 November 1990 (06.11.90) US

(71) Applicant: THE CHILDREN'S MEDICAL CENTER CORPORATION [US/US]; 55 Shattuck Street, Boston, MA 02115 (US).

(72) Inventor: EZEKOWITZ, Raymond, Alan, Brian ; 180 Beacon Street, Boston, MA 02116 (US).

(74) Agent: FREEMAN, John, W.; Fish & Richardson, 225 Franklin Street, 32nd Floor, Boston, MA 02110-2804 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

**Published**
*With international search report.*

(54) Title: SOLUBLE MANNOSE RECEPTOR PEPTIDES

(57) Abstract

Purified soluble recombinant peptides derived from an extracellular portion of the mannose receptor protein and fragments thereof, containing one or more carbohydrate recognition domains; nucleic acid producing these fragments, and vectors and cells including such nucleic acid are disclosed. The peptides are useful for treatment of disease.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| AT | Austria | ES | Spain | MG | Madagascar |
| AU | Australia | FI | Finland | ML | Mali |
| BB | Barbados | FR | France | MN | Mongolia |
| BE | Belgium | GA | Gabon | MR | Mauritania |
| BF | Burkina Faso | GB | United Kingdom | MW | Malawi |
| BG | Bulgaria | GN | Guinea | NL | Netherlands |
| BJ | Benin | GR | Greece | NO | Norway |
| BR | Brazil | HU | Hungary | PL | Poland |
| CA | Canada | IT | Italy | RO | Romania |
| CF | Central African Republic | JP | Japan | SD | Sudan |
| CG | Congo | KP | Democratic People's Republic of Korea | SE | Sweden |
| CH | Switzerland | | | SN | Senegal |
| CI | Côte d'Ivoire | KR | Republic of Korea | SU+ | Soviet Union |
| CM | Cameroon | LI | Liechtenstein | TD | Chad |
| CS | Czechoslovakia | LK | Sri Lanka | TG | Togo |
| DE* | Germany | LU | Luxembourg | US | United States of America |
| DK | Denmark | MC | Monaco | | |

\+ **Any designation of "SU" has effect in the Russian Federation. It is not yet known whether any such designation has effect in other States of the former Soviet Union.**

# SOLUBLE MANNOSE RECEPTOR PEPTIDES

## Background of the Invention

This invention relates to the general field of anti-microbial and anti-viral compounds, including diagnostic compounds, as well as to methods and reagents for making and using the compounds.

It is known to use as anti-microbial agents compounds that interfere with the metabolic processes of the infective cell. For example, antibacterial agents of the sulfonamide class, as structural analogs of p-aminobenzoic acid, block purine nucleotide synthesis in susceptible microorganisms, while penicillin prevents the completion of the final stages of cell wall biosynthesis.

A number of antiviral agents such as AZT and suramin achieve their effect by targeting the uniquely retroviral enzyme reverse transcriptase. AZT has been approved for treatment of patients with the Acquired Immune Deficiency Syndrome (AIDS), caused by the Human Immunodeficiency Virus Type 1 (HIV-1). Another anti-viral agent, the polyanionic compound dextran sulfate, blocks binding of virions to target cells. The soluble mannose-binding protein prevents infection of H9 lymphoblasts by HIV-1 by binding to the high mannose glycans expressed on the envelope glycoprotein of the retrovirus (Ezekowitz et al., J. Exp. Med. 169:185-196, 1989).

## Summary

In one aspect, the invention features a soluble recombinant peptide comprising at least one (and preferably two, three or more) carbohydrate recognition domains derived from an extracellular portion of mannose receptor protein (MRP). The peptide is capable of specifically targeting cells expressing mannose, N-

acetylglucosamine, or fucose, by virtue of those carbohydrate recognition domain(s) (CRD). For example, the MRP-derived carbohydrate recognition domain can specifically bind eucaryotic or procaryotic pathogenic cells (e.g., bacteria, fungi, or viruses) having exposed configurations of the specified sugar moieties on their cell wall or on the envelope glycoprotein. In addition, a peptide containing the MRP-derived CRD can specifically target cancer cells which have any exposed mannose residues as a result of abberant glycosylation. Peptides according to the invention offer a probe for such cells, or a tool for delivery of specific molecules (e.g., toxins or cell specific molecules such as the T-cell antigen, CD4) to those cells, or an in vivo marker for those cells to the immune system. The domain(s) are said to be MRP-derived in that they generally contain at least 150, or preferably 300 contiguous amino acids homologous to a sequence of one or more carbohydrate recognition domains of the mannose receptor protein, shown in Fig. 3. The soluble peptide lacks the transmembrane and cytoplasmic regions of MRP. By peptide is meant a chain of about ten or more amino acids, including larger polypeptides and proteins, that are useful in this invention. The peptide may be glycosylated via O- or N-linkages. By recombinant peptide is meant a peptide that is expressed from engineered nucleic acid, defined below.

In another aspect, the invention features engineered nucleic acid (preferably cDNA) encoding such a soluble peptide. By engineered nucleic acid is meant nucleic acid removed from its natural environment (i.e., from naturally adjacent nucleic acid) by purification or recombinant DNA methodology; the term also includes synthetic nucleic acid or cDNA. This nucleic acid may be a fragment of DNA or RNA, it may be present in a vector system (e.g., a plasmid, cosmid or phage), or it may be

within the genome of an organism. In some cases, such nucleic acid is purified and includes a homogeneous preparation of desired nucleic acid.

In preferred embodiments, the peptide and nucleic acid encoding it are further characterized by at least 75% identity at the amino acid level to a sequence of at least 150, and preferably 300, contiguous amino acids of one or more carbohydrate recognition domains of mannose receptor protein most preferably the peptide includes the entire extracellular region of mannose receptor protein. In other preferred embodiments, the nucleic acid substantially corresponds to at least 450 contiguous bases of the nucleic acid encoding the soluble extracellular fragment of mannose receptor protein, deposited in the ATCC as ATCC No. 68430 and described herein as nucleotides 1-4212 of SEQ ID NO: 1; and the nucleic acid is ligated to nucleic acid encoding the toxic part of a toxin molecule (e.g., AZT, ricin, or cholera toxin), or to nucleic acid encoding a peptide capable of fixing complement. The hybrid peptides encoded by such ligated nucleic acid are especially useful for causing an effector molecule to be targeted to an undesired cell or other organism, such as a virus.

The peptides described above, and antibodies to those peptides, may be used in therapeutic or diagnostic agents. Preferably the peptide is purified, that is, the peptide is substantially separated from contaminating peptides. Most preferably it is provided as a homogenous preparation admixed in a carrier substance suitable for therapeutic use. By therapeutic agent is meant a substance useful for the treatment of a disease or disorder; by diagnostic agent is meant a substance relating to the detection of a disease or disorder.

In yet other aspects, the invention features methods for treating an animal, e.g., a human, infected

with a bacterium, fungus, or virus. (By bacterium, fungus, or virus is meant to include any type of undesired cell or other organism that is capable of causing an infection.) One such method includes providing and administering a therapeutically effective amount of a therapeutic agent or peptide including a soluble extracellular portion of mannose receptor protein capable of specifically targeting cells expressing mannose, N-acetylglucosamine, or fucose. The therapeutic agent or peptide causes direct inhibition of growth of the infective organism, or causes host defensive cells, e.g., macrophages, to be attracted to the pathogenic organisms which are thereby inactivated. Such inactivation may be aided by the presence of complement which is fixed by the peptide. A therapeutically effective amount is that quantity which produces a significant physiological effect in the patient and is recognized by those of ordinary skill in the art to depend upon the size and weight of the animal as well as other well known factors.

In preferred embodiments, the peptide is a therapeutically effective fragment of the soluble extracellular portion of mannose receptor protein; the peptide is able to inhibit (e.g., reduce or prevent) growth of, or infection by, the bacterium, fungus, or virus, and is a peptide as described above. Most preferably, the animal is human; the infection is one that results in a bacteremia or local bacterial infection, parasitic infection, or fungal colonization, and the route of administration is either intravenous, intramuscular, oral, or local, e.g., in the form of a powder, or lotion, preferably at 5-100 μg/ml, more preferably at 25 μg/ml; or the virus is HIV or a related virus, and the peptide lowers the rate of infection of eucaryotic cells by the virus; the protein or peptide is

provided at 1-500 μg/ml (preferably 100-150 μg/ml) final concentration in human serum or tissue. Alternatively, lipid vesicles, or lyposomes, containing toxins or antibiotics are coated with the peptide and administered directly to the patient. Such lyposomes will be targeted to the infected area by the peptide and the content of the lyposomes released, thereby specifically retarding or preventing growth of the targeted cells or organisms in the targeted area.

In a related aspect, the invention features a coated catheter, useful for long-term administration of fluids to a patient. The catheter is coated with one of the above-described peptides, e.g., by impregnating the catheter material with the peptide. The peptide lowers the rate of bacterial, fungal or viral infection of the patient through the catheter.

In another aspect, the invention features a method for diagnosing infection by a bacterium, fungus or virus. The method includes detecting the serum level of a pathogen that expresses one of the target glycoproteins recognized by MRP, by measuring the amount of binding of a peptide according to the invention to a sample of the serum. The detected pathogen level reflects the infection of the patient. Preferably, the method features measuring the peptide by immunologic or fluorescent techniques.

In a related aspect, the invention features a purified antibody specifically recognizing a peptide according to the invention. The antibody is preferably provided as a homogeneous preparation of a monoclonal or polyclonal antibody. The antibody is useful for purification of the extracellular portion of mannose receptor protein or peptides thereof, according to the invention, and for diagnosis of infection as disclosed above.

In a final aspect, the invention features a purified soluble peptide comprising the extracellular portion of mannose receptor protein, said peptide lacking the mannose receptor protein transmembrane and cytoplasmic regions.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

Brief Description of the Drawings

Fig. 1a shows the nucleotide base sequence and the corresponding amino acid sequence of the extracellular portion of the mannose receptor protein, described herein as nucleotides 1-4212 of SEQ ID NO: 1.

Fig. 1b shows the nucleotide base sequence and the corresponding amino acid sequence of the transmembrane and cytoplasmic portions of the mannose receptor protein, described herein as nucleotides 1-4212 of SEQ ID NO: 2.

Fig. 2 is a schematic diagram including the functional regions of the extracellular portion of the mannose receptor protein.

Fig. 3 shows the correspondance of the amino acid sequence of the various carbohydrate recognition domains of the mannose receptor, using the single letter amino acid code.

Description of the Preferred Embodiment

Soluble recombinant peptides derived from an extracellular portion of the mannose receptor that contains one or more carbohydrate recognition domains (CRDs) are able to recognize carbohydrates with a specificity comparable to that of the native membrane-bound mannose receptor. Such soluble mannose receptor peptides can be immobilized or attached to a portion of another molecule without losing effective carbohydrate recognition capacity.

Exposed sugars like mannose and N-acetylglucosamine are a feature of the cell walls of many pathogens, whereas higher organisms, including humans and animals, tend to have masked internal mannose residues that are not recognized by the mannose receptor. Therefore, soluble mannose receptor peptides according to the invention are useful in therapeutic agents in that they specifically bind mannose-rich pathogens, including bacteria, fungi, yeasts, parasites, or the envelope glycoproteins of certain viruses. Such peptides can also specifically target cancer cells having exposed mannose residues as a result of abberant glycosylation. When such soluble peptides are attached to other entities such as macrophages or peptide portions that fix complement, or used as a tool for the delivery of specific molecules such as toxins or cell-specific agents to mannose-rich pathogens, the peptides can direct removal of such pathogens from the patient. The soluble peptides according to the invention are also useful as probes in diagnosis.

The amino acid sequence of the mannose receptor, from which peptides according to the invention are derived, is inferred from the cDNA sequence (Figs. 1a and 1b) and analyzed below:

Referring to Fig. 2, the first domain is comprised of 134 amino acids at the $NH_2$ terminus. Without being bound to any theory, it appears that this cysteine-rich region is not essential to mannose targeting according to the invention, and, therefore, it may be deleted without departing from the spirit of the invention. Preferred soluble peptides according to the invention, however, include this domain.

The second domain spans from residues 135-188. Without being bound to any theory, this domain appears to be related to fibronectin type II, and it may play a role

in interaction with the extracellular matrix and contribute to the spreading and adhesion of tissue macrophages expressing the full length receptor. As with the first domain, the second domain is not essential to the practice of the invention, but preferred peptides include it.

Carbohydrate recognition domains (CRDs) comprise the remainder of the extracellular portion of the receptor. Specifically, there are eight segments related to C-type carbohydrate recognition domains (CRDs) of animal lectins reported by Drickamer, J. Biol. Chem. 263:9557-9560 (1988). These CRDs are discussed in greater detail below, and they are central to the invention.

A transmembrane region and a COOH-terminal cytoplasmic domain are truncated from the receptor in peptides according to the invention to enhance solubility and facilitate therapeutic application of such peptides. Surprisingly, after truncation of the transmembrane and intracellular portions of the receptor, the molecule retains carbohydrate binding capacity effective for various purposes discussed elsewhere in this application.

Preferred peptides according to the invention include at least one or more, and preferably four or more, of the eight carbohydrate recognition domains (CRDs) depicted in Fig. 2. The sequences of the individual CRDs are shown in Fig. 3 with the numbering of the starting amino acid of each CRD keyed to the amino acid sequence of the entire soluble extracellular portion of the mannose receptor as shown in Fig. 1a, nucleotides 1-4212 of SEQ ID NO: 1.

Those skilled in the art will recognize that it is possible to vary the specific sequence of the soluble carbohydrate-targeting peptide being used, without deviating from the concept and spirit of the invention.

The entire extracellular portion of the mannose receptor, truncated to remove the transmembrane portion and the cytoplasmic tail of the full length receptor, is a preferred peptide, but CRD-containing fragments of the truncated receptor are also within the scope of the invention.

Not only does the invention cover CRD containing fragments of the truncated receptor, it also covers conservative mutations of the truncated receptor and its fragments. Preferably, a peptide according to the invention includes multiple (two or more and, preferably, four or more) CRDs of mannose receptor protein. Moreover, individual CRDs can be repeated to further increase the carbohydrate binding capacity of the peptide according to the invention. Merely by way of example, and not as a limitation, the peptide can include multiple copies of one or more of the specific CRDs of the mannose receptor, shown in Fig. 3.

Isolated nucleic acid encoding a CRD of mannose receptor protein is useful for producing recombinant peptide fragments of the protein. In addition, the nucleic acid can be modified by standard techniques in order to express the same or modified peptides; e.g., by conservative base substitution the nucleic acid can be modified and still encode the same amino acid sequence, or the nucleic acid can be modified to encode a conservative amino acid substitution, which will preserve the tertiary structure and the distribution of charged amino acids in the peptide.

We now describe a specific cDNA clone of the extracellular portion (ectodomain) of mannose receptor protein (MRP). The clone is described not only as a specific example of the invention but also as a starting material to obtain other peptides according to the invention, using methods of producing candidate peptides

and methods for screening such candidates for mannose affinity, as described below.

Example 1: Cloning of Full Length MRP and Processing of MRP cDNA to encode and express peptides of the invention

Sequences for probes were determined by obtaining sequence information from purified receptor. Receptor was purified from alveolar macrophages or human placenta as described (Lennartz et al., J. Biol. Chem. 262:9942, 1987).

Degenerate oligonucleotide probes were synthesized on a Dupont oligonucleotide synthesizer, purified by gel filtration, and labeled with $^{32}$p-ATP and polynucleotide kinase. Radiolabled probe was used to screen a pCDM8 placental cDNA library (gift of Dr. B. Seed, Harvard Medical School) by colony hybridization. Twenty-five positive clones were isolated by two rounds of amplification and analyzed. The longest clone (3.3kb) was found upon analysis to contain sequences encoding a number of peptides that had been derived from the placental mannose receptor (Taylor et al., J. Biol. Chem. 265:12156, 1990). This 3.3kb placental derived clone was radiolabeled and used as a probe to isolate the macrophage mannose receptor cDNAs from a 7 day macrophage cDNA library (Ezekowitz et al., J. Exp. Med., in press, December, 1990). A 750bp cDNA derived from the 5' extent of the placental mannose receptor cDNA was utilized to isolate 5' clones from the macrophage library. A full-length cDNA was then assembled in a CDM8 expression vector.

The sequence of macrophage mannose receptor is identical to the placental form except for a C to T polymorphism at nucleotide 2284. The initial placental clone was sequenced by double stranded sequencing using a modified T7 polymerase, Sequenase® (U.S. Biochemical, Cleveland, Ohio) based on the Sanger chain termination

method (Sanger et al., Proc. Natl. Acad. Sci., USA 74:5463, 1977). Specific oligonucleotides were synthesized and used as sequencing primers. For phage clones, 2μl of purified stock was annealed to λGT11 primers from each of the arms and the taq polymerase amplified product obtained after 25 cycles (94°C, 30s denaturation, 55°C 30s annealing, and 72°C 3 minute extension) on a thermal cycler (Dorfman et al., Bio. Techniques, 7:568, 1989), and the products were gel purified by agarose gel electrophoreisis. The purified products were digested with EcoRI, subcloned into a pUC-19 vector, and the nucleotide sequence determined as described above.

The encoded protein sequence deduced from the nucleotide sequence is shown in Figs. 1a and 1b (SEQ ID NO: 1). The open reading frame predicts a protein of 1438 amino acids which is consistent with the estimated molecular weight of the receptor polypeptide (150kD) after the N-linked sugars have been removed. (Lennartz et al., J. Biol. Chem. 264:2385, 1989; Taylor et al., J. Biol. Chem. 265:12156, 1990; Ezekowitz et al., J. Exp. Med., in press, December, 1990).

The features of the membrane bound mannose receptor protein are depicted in a schematic diagram (Fig. 2) and include (i) a typical hydrophobic signal peptide; (ii) a cysteine rich $NH_2$ terminal region; (iii) a fibronectin type II domain; (iv) eight carbohydrate recognition domains; (v) a hydrophobic transmembrane region; and (vi) a cytoplasmic tail. The $NH_2$ terminal amino acid is defined by an N-terminus peptide as Leu, which is preceded by Ala-Val-Leu, a typical recognition sequence for a signal peptidase (Von Heijne, Eur. J. Biochem. 133:17, 1983).

For encoding and expressing peptides of the invention, cDNA encoding the full length mannose receptor

protein (MRP) was first derived in a CDM8 plasmid expression vector as described above. A construct of the cDNA encoding soluble mannose receptor peptide was then prepared in a CDM8 plasmid by a multiple step procedure, as follows.

Referring to Figs. 1a and 1b, in the first step: an antisense primer was designed from a 3' end, at base pair 4169, to a 5' end, at base pair 4201, to contain a HpaI site. The sense primer was prepared from a 5' end at base pair 3475 and encompassed the NsiI site base pair 3510. The primers were annealed to full length mannose receptor cDNA, and a 726 base pair fragment was amplified using the polymerase chain reaction technique (PCR). The full length cDNA mannose receptor in CDM8 was then digested with NsiI and HpaI which released a fragment from the unique NsiI site in the cDNA to the HpaI site in the vector, thereby removing (see Figs. 1a and 2) the last three amino acids of the ectodomain, the entire transmembrane region, the entire cytoplasmic domain, and some vector sequence. This fragment was replaced with the 726 bp PCR fragment, thereby creating a clone (SMR), confirmed by sequence analysis, which contained cDNA encoding the signal peptide and the entire ectodomain of the mannose receptor (except for the last three amino acids). This clone is capable of generating a soluble mannose receptor peptide. This construct can be transfected stably or transiently into a mammalian expression system, and the soluble receptor peptide expressed is secreted into the medium.

From this plasmid, which has been deposited in the ATCC as ATCC No. 68430, a series of truncated forms of soluble mannose receptor peptide containing various numbers of carbohydrate recognition domains can be constructed by standard molecular biological techniques, either by using the polymerase chain reaction or

convenient restriction enzyme sites to create molecules that can be secreted. Alternatively, standard molecular biological techniques may be used to isolate other nucleic acid (especially cDNA) clones encoding the extracellular portion of the mannose receptor protein by procedures analogous to those described above.

Expression vectors suitable for peptide expression also include standard bacterial, yeast, and viral expression vectors, as well as eucaryotic vectors. Those skilled in the art will realize that such vectors generally are suitable for expressing peptides of the invention.

Expression of soluble human mannose receptor peptides by these vectors and organisms can be followed using a mannan affinity column such as sepharose-mannose. The column is first contacted with the expressed material. Peptides able to recognize and bind mannose are bound to the mannose-sepharose matrix, eluted with 50mM Tris/10M EDTA, and identified using 8% polyacrylamide gels (with Laemmli buffers, Nature 227:600, 1970). Those clones which produce peptides able to bind to such a column are among those useful in this invention.

USE OF THE PEPTIDES

Soluble mannose receptor peptides expressed as described above are useful for specifically targeting (or specifically recognizing) cells expressing carbohydrates such as mannose, N-acetylglucosamine, or fucose on their surface. Thus these peptides are useful in agents for diagnosing or treating infection by a wide variety of pathogenic organisms, e.g., Leishmania proamastigotes, Pneumocystis carinii, Candida albicans, Microbacteria tuberculosis (and other atypical mycobacteria), Human Immunodeficiency Virus Type 1 (HIV-1) or influenza virus. Such agents are also useful for treating opportunistic

infections such as those that arise in patients with cancer, patients undergoing chemotherapy or bone marrow transplants, or patients suffering from congenital or acquired immune deficiency diseases, such as AIDS. In addition, such agents can specifically target cancer cells having exposed mannose residues as a result of abberant glycosylation.

For non-viral pathogens, removal by host defense mechanisms is achieved by directing attachment of a soluble mannose receptor peptide, in conjunction with the cell attachment site of a receptor such as the mannose-binding protein, to the surface of phagocytic cells, thereby enhancing the clearance of the pathogens from the circulation by causing the phagocytes to recognize the pathogens. For viruses which express mannose-rich glycoproteins, direct inactivation of the virus and viral infected cells is accomplished by attaching toxins, such as ricin, cholera, diphtheria, or pertussis, or antimetabolic drugs, such as AZT, to a therapeutic soluble mannose receptor peptide. The hybrid peptide thus formed can serve to kill or inhibit growth of the target cell, such as HIV.

To form such hybrid peptides, nucleic acid encoding such toxins can be ligated by well known techniques to nucleic acid encoding a soluble mannose receptor peptide according to the invention, and the fused nucleic acid can be expressed as a single entity to form a hybrid peptide (for example, as described by Murphy, U.S. Patent No. 4,675,383, hereby incorporated by reference). (By ligated is meant linked enzymatically or chemically to form a single nucleic acid entity.) Alternatively, the two peptides can be synthesized separately and linked chemically (for example, as described by Ross, U.S. Patent No. 4,275,000, hereby incorporated by reference).

Alternatively, nucleic acid encoding a complement-fixing region, e.g., the complement-fixing region of the immunoglobulin heavy chain or of the mannose-binding protein, can be engineered by standard techniques to form a hybrid molecule with nucleic acid encoding a soluble mannose receptor protein. The expression product of such nucleic acid can be used to target cells with exposed surface carbohydrate moieties and then to interact with complement components and activate complement. Activated complement will then stimulate binding of macrophages to the targeted pathogenic cells and their subsequent ingestion by the macrophages.

Example 2: Preparation of a fusion protein

A soluble mannose receptor peptide-immunoglobulin fusion protein can be prepared by digesting cDNA encoding soluble mannose receptor peptide and inserting an oligonucleotide linker (e.g., a BamH1 linker). The resulting plasmid can be digested with BamH1, and the portion encoding the entire extracellular domain can be ligated to the synthetic splice donor sequence of an immunoglobulin (e.g., human IgG1) expression plasmid (Aruffo et al., Cell 61:1303-1313, 1990). Such expression vectors contain in their 3' region the immunoglobulin heavy chain constant regions two and three, which have the capacity to fix complement. A fusion protein expressed from such a fused cDNA sequence would contain a complement-fixing region at the 3' end of a soluble mannose receptor peptide.

In another construct, cDNA encoding an immunoglobulin signal peptide fused to the $NH_2$ terminal region, cell-attachment domain, and complement-fixing region of the mannose-binding protein (Ezekowitz, International Patent Application No. WO 89/01519, February 23, 1989) can be engineered to replace the

cysteine-rich and fibronectin binding domains of a soluble mannose receptor peptide. The fusion protein expressed from such a cDNA sequence would contain a complement-fixing region in the amino terminal portion of the molecule, preceding the carbohydrate recognition domains.

Soluble mannose receptor peptides according to the invention may be administered by routine methods in pharmaceutically acceptable carrier substances, i.e., inert substances suitable for pharmaceutical use such as the dispensing of drugs or medicine. For example, they can be administered in an aerosol form to treat, e.g., Pneumocystis carinii. Alternatively, they may be administered orally or parenterally, e.g., they can be injected directly into the blood stream of an animal, especially humans, to a level of between 1-500 μg/ml serum (most preferably, 100-150 μg/ml) final concentration, and this dose repeated to maintain this level. The peptides can be administered prophylactically or after infection.

In a specific prophylactic use, soluble mannose receptor peptides may be used to coat intravenous or urethral catheters (e.g., by chemical impregnation of the catheter material with the peptide) to prevent infection in immunocompromised patients (e.g., cancer patients subjected to long term intravenous chemotherapy). Such catheters will bind infective organisms and prevent their entry into the patient.

In a specific therapeutic use, the soluble mannose receptor peptide may be applied topically in powder or lotion form (at a concentration of between 56-100 μg/ml), for example, to treat local infections, such as bacterial infection, yeast infection, or infection with Trichophyton, which causes athlete's foot.

Soluble mannose receptor peptides can also be used as a diagnostic tool, e.g., for the diagnosis of fungal diseases. Fungi infecting an animal will shed a mannose-rich polysaccharide into the serum. A sample of serum from a patient (e.g., 100μl) can be analyzed with fluorescently-labelled soluble mannose receptor peptide to observe binding to the fungal polysaccharide coat, and the degree of binding can be related to the degree of fugal infection remaining following a course of treatment. In an alternative diagnostic method, the degree of binding of the soluble mannose receptor peptide can be detected by using a labelled antibody that specifically recognizes the peptide. Appropriate subsequent treatment can be planned accordingly.

Other Embodiments

As described above, the invention generally features peptides that include a MRP-derived carbohydrate recognition domain. The genetic material encoding soluble mannose receptor peptide (deposited as described above as ATCC No. 68430) can be used to generate a large number of recombinant peptides by fragmenting the full-length nucleic acid and expressing candidate fragments. Alternatively, as described above, standard molecular biological techniques may be used to isolate other nucleic acid (especially cDNA) clones encoding soluble mannose receptor peptides. These clones can also be used to express candidate fragment peptides. As described, preferred fragments are those containing multiple CRDs. Various assays may be used to determine whether a particular candidate peptide has carbohydrate recognition ability.

In addition to the affinity column chromatographic assay described above, another assay invokes binding and uptake of $I^{125}$-labeled mannose-BSA. Specifically, mannose-BSA (EY Labs, CA) is radiolabeled as described

(Ezekowitz et al., J. Exp. Med. 154:60, 1981), and binding and uptake of radiolabled ligand is performed on COS-I cells transfected with cloned cDNA encoding the candidate peptide. COS-I cells transfected with CD64 serve as controls, and thioglycollate elicited mouse peritoneal macrophages serve as a positive control.

Another assay utilizes antibody that specifically recognizes a soluble mannose receptor peptide. The antibody may be linked with a fluorescent tag and antibody-peptide binding identified flow cytometrically. Alternatively, the antibody may be immobilized for assay use or employed in an enzyme linked immunosorbent assay or ELIZA test.

Deposits

Plasmid SMR, in CDM8 in E-coli strain MC1061/P3, was deposited on Oct. 2, 1990, with the American Type Culture Collection (ATCC) as ATCC No. 68430.

Applicant's assignee, Children's Medical Center Corporation, represents that the ATCC is a depository affording permanence of the deposit and ready accessibility thereto by the public if a patent is granted. All restrictions on the availability to the public of the material so deposited will be irrevocably removed upon the granting of a patent. The material will be available during the pendency of the patent application to one determined by the Commissioner to be entitled thereto under 37 CFR 1.14 and 35 USC 122. The deposited material will be maintained with all the care necessary to keep it viable and uncontaminated for a period of at least five years after the most recent request for the furnishing of a sample of the deposited microorganism, and in any case, for a period of at least thirty (30) years after the date of deposit or for the enforceable life of the patent, whichever period is longer. Applicant's assignee acknowledges its duty to

replace the deposit should the depository be unable to furnish a sample when requested due to the condition of the deposit.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Ezekowitz, Raymond Alan Brian

(ii) TITLE OF INVENTION: SOLUBLE MANNOSE RECEPTOR PEPTIDES

(iii) NUMBER OF SEQUENCES: 1

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Fish & Richardson
(B) STREET: One Financial Center
(C) CITY: Boston
(D) STATE: Massachusetts
(E) COUNTRY: U.S.A.
(F) ZIP CODE: 02111-2658

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
(B) COMPUTER: IBM PS/2 Model 50Z or 55SX
(C) OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
(D) SOFTWARE: WordPerfect (Version 5.0)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

Prior applications total, including application described below: 0

(A) APPLICATION NUMBER: 0
(B) FILING DATE: 0

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Freeman, John W.
(B) REGISTRATION NUMBER: 29,066
(C) REFERENCE/DOCKET NUMBER: 00108-032001

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 542-5070
(B) TELEFAX: (617) 542-8906
(C) TELEX: 200154

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 1

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5145 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO. 1:

```
TTGCGTCTTA GTTCCGCCCT CCTGTCCATC AGGAGAAGGA AAGGATAAAC CCTGGGCC        63

ATG AGG CTA CCC CTG CTC CTG GTT TTT GCC TCT GTC ATT CCG GGT GCT        111
Met Arg Leu Pro Leu Leu Leu Val Phe Ala Ser Val Ile Pro Gly Ala
           -15                 -10                 -5

GTT CTC CTA CTG GAC ACC AGG CAA TTT TTA ATC TAT AAT GAA GAT CAC        159
Val Leu Leu Leu Asp Thr Arg Gln Phe Leu Ile Tyr Asn Glu Asp His
     1               5                  10

AAG CGC TGC GTG GAT GCA GTG AGT CCC AGT GCC GTC CAA ACC GCA GCT        207
Lys Arg Cys Val Asp Ala Val Ser Pro Ser Ala Val Gln Thr Ala Ala
15                  20                  25                  30

TGC AAC CAG GAT GCC GAA TCA CAG AAA TTC CGA TGG GTG TCC GAA TCT        255
Cys Asn Gln Asp Ala Glu Ser Gln Lys Phe Arg Trp Val Ser Glu Ser
                35                  40                  45

CAG ATT ATG AGT GTT GCA TTT AAA TTA TGC CTG GGA GTG CCA TCA AAA        303
Gln Ile Met Ser Val Ala Phe Lys Leu Cys Leu Gly Val Pro Ser Lys
           50                  55                  60

ACA GAC TGG GTT GCT ATC ACT CTC TAT GCC TGT GAC TCA AAA AGT GAA        351
Thr Asp Trp Val Ala Ile Thr Leu Tyr Ala Cys Asp Ser Lys Ser Glu
     65                  70                  75

TTT CAG AAA TGG GAG TGC AAA AAT GAC ACA CTT TTG GGG ATC AAA GGA        399
Phe Gln Lys Trp Glu Cys Lys Asn Asp Thr Leu Leu Gly Ile Lys Gly
80                  85                  90

GAA GAT TTA TTT TTT AAC TAC GGC AAC AGA CAA GAA AAG AAT ATT ATG        447
Glu Asp Leu Phe Phe Asn Tyr Gly Asn Arg Gln Glu Lys Asn Ile Met
95                  100                 105                 110

CTC TAC AAG GGA TCG GGT TTA TGG AGC AGG TGG AAG ATC TAT GGA ACC        495
Leu Tyr Lys Gly Ser Gly Leu Trp Ser Arg Trp Lys Ile Tyr Gly Thr
                115                 120                 125

ACA GAC AAT CTG TGC TCC AGA GGT TAT GAA GCC ATG TAT ACG CTA CTA        543
Thr Asp Asn Leu Cys Ser Arg Gly Tyr Glu Ala Met Tyr Thr Leu Leu
           130                 135                 140

GGC AAT GCC AAT GGA GCA ACC TGT GCA TTC CCG TTC AAG TTT GAA AAC        591
Gly Asn Ala Asn Gly Ala Thr Cys Ala Phe Pro Phe Lys Phe Glu Asn
     145                 150                 155

AAG TGG TAC GCA GAT TGC ACG AGT GCT GGG CGG TCG GAT GGA TGG CTC        639
Lys Trp Tyr Ala Asp Cys Thr Ser Ala Gly Arg Ser Asp Gly Trp Leu
160                 165                 170

TGG TGC GGA ACC ACT ACT GAC TAT GAC ACA GAC AAG CTA TTT GGA TAT        687
Trp Cys Gly Thr Thr Thr Asp Tyr Asp Thr Asp Lys Leu Phe Gly Tyr
175                 180                 185                 190

TGT CCA TTG AAA TTT GAG GGC AGT GAA AGC TTA TGG AAT AAA GAC CCG        735
Cys Pro Leu Lys Phe Glu Gly Ser Glu Ser Leu Trp Asn Lys Asp Pro
                195                 200                 205

CTG ACC AGC GTT TCC TAC CAG ATA AAC TCC AAA TCC GCT TTA ACG TGG        783
Leu Thr Ser Val Ser Tyr Gln Ile Asn Ser Lys Ser Ala Leu Thr Trp
           210                 215                 220

CAC CAA GCG AGG AAA AGC TGC CAA CAA CAG AAC GCT GAG CTC CTG AGC        831
```

```
His Gln Ala Arg Lys Ser Cys Gln Gln Gln Asn Ala Glu Leu Leu Ser
      225                 230                 235

ATC ACA GAG ATA CAT GAG CAA ACA TAC CTG ACA GGA TTA ACC AGT TCC      879
Ile Thr Glu Ile His Glu Gln Thr Tyr Leu Thr Gly Leu Thr Ser Ser
  240                 245                 250

TTG ACC TCA GGA CTC TGG ATT GGA CTT AAC AGT CTG AGC TTC AAC AGC      927
Leu Thr Ser Gly Leu Trp Ile Gly Leu Asn Ser Leu Ser Phe Asn Ser
255                 260                 265                 270

GGT TGG CAG TGG AGT GAC CGC AGT CCT TTC CGA TAT TTG AAC TGG TTA      975
Gly Trp Gln Trp Ser Asp Arg Ser Pro Phe Arg Tyr Leu Asn Trp Leu
                275                 280                 285

CCA GGA AGT CCA TCA GCT GAA CCT GGA AAA AGC TGT GTG TCA CTA AAT      1023
Pro Gly Ser Pro Ser Ala Glu Pro Gly Lys Ser Cys Val Ser Leu Asn
            290                 295                 300

CCT GGA AAA AAT GCT AAA TGG GAA AAT CTG GAA TGT GTT CAG AAA CTG      1071
Pro Gly Lys Asn Ala Lys Trp Glu Asn Leu Glu Cys Val Gln Lys Leu
      305                 310                 315

GGC TAT ATT TGC AAA AAG GGC AAC ACC ACT TTA AAT TCT TTT GTT ATT      1119
Gly Tyr Ile Cys Lys Lys Gly Asn Thr Thr Leu Asn Ser Phe Val Ile
  320                 325                 330

CCC TCA GAA AGT GAT GTG CCT ACT CAC TGT CCT AGT CAG TGG TGG CCG      1167
Pro Ser Glu Ser Asp Val Pro Thr His Cys Pro Ser Gln Trp Trp Pro
335                 340                 345                 350

TAT GCC GGT CAC TGT TAC AAG ATT CAC AGA GAT GAG AAA AAA ATC CAG      1215
Tyr Ala Gly His Cys Tyr Lys Ile His Arg Asp Glu Lys Lys Ile Gln
                355                 360                 365

AGG GAT GCT CTG ACC ACC TGC AGG AAG GAA GGC GGT GAC CTC ACA AGT      1263
Arg Asp Ala Leu Thr Thr Cys Arg Lys Glu Gly Gly Asp Leu Thr Ser
            370                 375                 380

ATC CAC ACC ATC GAG GAA TTG GAC TTT ATT ATC TCC CAG CTA GGA TAT      1311
Ile His Thr Ile Glu Glu Leu Asp Phe Ile Ile Ser Gln Leu Gly Tyr
      385                 390                 395

GAG CCA AAT GAC GAA TTG TGG ATC GGC TTA AAT GAC ATT AAG ATT CAA      1359
Glu Pro Asn Asp Glu Leu Trp Ile Gly Leu Asn Asp Ile Lys Ile Gln
  400                 405                 410

ATG TAC TTT GAG TGG AGT GAT GGG ACC CCT GTA ACG TTT ACC AAA TGG      1407
Met Tyr Phe Glu Trp Ser Asp Gly Thr Pro Val Thr Phe Thr Lys Trp
415                 420                 425                 430

CTT CGT GGA GAA CCA AGC CAT GAA AAC AAC AGA CAG GAG GAT TGT GTG      1455
Leu Arg Gly Glu Pro Ser His Glu Asn Asn Arg Gln Glu Asp Cys Val
                435                 440                 445

GTG ATG AAA GGC AAG GAT GGG TAC TGG GCA GAT CGG GGC TGT GAG TGG      1503
Val Met Lys Gly Lys Asp Gly Tyr Trp Ala Asp Arg Gly Cys Glu Trp
            450                 455                 460

CCT CTT GGC TAC ATC TGC AAG ATG AAA TCA CGA AGC CAA GGT CCA GAA      1551
Pro Leu Gly Tyr Ile Cys Lys Met Lys Ser Arg Ser Gln Gly Pro Glu
            465                 470                 475
```

```
ATA GTG GAA GTC GAA AAA GGC TGC AGG AAA GGC TGG AAA AAA CAT CAC      1599
Ile Val Glu Val Glu Lys Gly Cys Arg Lys Gly Trp Lys Lys His His
        480                 485                 490

TTT TAC TGC TAT ATG ATT GGA CAT ACG CTT TCA ACA TTT GCA GAA GCA      1647
Phe Tyr Cys Tyr Met Ile Gly His Thr Leu Ser Thr Phe Ala Glu Ala
    495                 500                 505

AAC CAA ACC TGT AAT AAT GAG AAT GCT TAT TTA ACA ACT ATT GAA GAC      1695
Asn Gln Thr Cys Asn Asn Glu Asn Ala Tyr Leu Thr Thr Ile Glu Asp
510                 515                 520                 525

AGA TAT GAA CAA GCC TTC CTG ACT AGT TTC GTT GGC TTA AGG CCT GAA      1743
Arg Tyr Glu Gln Ala Phe Leu Thr Ser Phe Val Gly Leu Arg Pro Glu
                530                 535                 540

AAA TAT TTC TGG ACA GGA CTT TCA GAT ATA CAA ACC AAA GGG ACT TTT      1791
Lys Tyr Phe Trp Thr Gly Leu Ser Asp Ile Gln Thr Lys Gly Thr Phe
            545                 550                 555

CAG TGG ACC ATC GAG GAA GAG GTT CGG TTC ACC CAC TGG AAT TCA GAT      1839
Gln Trp Thr Ile Glu Glu Glu Val Arg Phe Thr His Trp Asn Ser Asp
        560                 565                 570

ATG CCA GGG CGA AAG CCA GGG TGT GTT GCC ATG AGA ACC GGG ATT GCA      1887
Met Pro Gly Arg Lys Pro Gly Cys Val Ala Met Arg Thr Gly Ile Ala
    575                 580                 585

GGG GGC TTA TGG GAT GTT TTG AAA TGT GAT GAA AAG GCA AAA TTT GTG      1935
Gly Gly Leu Trp Asp Val Leu Lys Cys Asp Glu Lys Ala Lys Phe Val
590                 595                 600                 605

TGC AAG CAC TGG GCA GAA GGA GTA ACC CAC CCA CCG AAG CCC ACG ACG      1983
Cys Lys His Trp Ala Glu Gly Val Thr His Pro Pro Lys Pro Thr Thr
                610                 615                 620

ACT CCC GAA CCC AAA TGT CCG GAG GAT TGG GGC GCC AGC AGT AGA ACA      2031
Thr Pro Glu Pro Lys Cys Pro Glu Asp Trp Gly Ala Ser Ser Arg Thr
            625                 630                 635

AGC TTG TGT TTC AAG CTG TAT GCA AAA GGA AAA CAT GAG AAG AAA ACG      2079
Ser Leu Cys Phe Lys Leu Tyr Ala Lys Gly Lys His Glu Lys Lys Thr
        640                 645                 650

TGG TTT GAA TCT CGA GAT TTT TGT CGA GCT CTG GGT GGA GAC TTA GCT      2127
Trp Phe Glu Ser Arg Asp Phe Cys Arg Ala Leu Gly Gly Asp Leu Ala
    655                 660                 665

AGC ATC AAT AAC AAA GAG GAA CAG CAA ACA ATA TGG CGA TTA ATA ACA      2175
Ser Ile Asn Asn Lys Glu Glu Gln Gln Thr Ile Trp Arg Leu Ile Thr
670                 675                 680                 685

GCT AGT GGA AGC TAC CAC AAA CTG TTT TGG TTG GGA TTG ACA TAT GGA      2223
Ala Ser Gly Ser Tyr His Lys Leu Phe Trp Leu Gly Leu Thr Tyr Gly
                690                 695                 700

AGC CCT TCA GAA GGT TTT ACT TGG AGT GAT GGT TCT CCT GTT TCA TAT      2271
Ser Pro Ser Glu Gly Phe Thr Trp Ser Asp Gly Ser Pro Val Ser Tyr
            705                 710                 715

GAA AAC TGG GCT TAT GGA GAA CCT AAT AAT TAT CAA AAT GTT GAA TAC      2319
Glu Asn Trp Ala Tyr Gly Glu Pro Asn Asn Tyr Gln Asn Val Glu Tyr
        720                 725                 730
```

```
TGT GGT GAG CTG AAA GGT GAC CCT ACT ATG TCT TGG AAT GAT ATT AAT      2367
Cys Gly Glu Leu Lys Gly Asp Pro Thr Met Ser Trp Asn Asp Ile Asn
    735                 740                 745

TGT GAA CAC CTT AAC AAC TGG ATT TGC CAG ATA CAA AAA GGA CAA ACA      2415
Cys Glu His Leu Asn Asn Trp Ile Cys Gln Ile Gln Lys Gly Gln Thr
750                 755                 760                 765

CCA AAA CCT GAG CCA ACA CCA GCT CCT CAA GAC AAT CCA CCA GTT ACT      2463
Pro Lys Pro Glu Pro Thr Pro Ala Pro Gln Asp Asn Pro Pro Val Thr
                770                 775                 780

GAA GAT GGG TGG GTT ATT TAC AAA GAC TAC CAG TAT TAT TTC AGC AAA      2511
Glu Asp Gly Trp Val Ile Tyr Lys Asp Tyr Gln Tyr Tyr Phe Ser Lys
            785                 790                 795

GAG AAG GAA ACC ATG GAC AAT GCG CGA GCG TTT TGC AAG AGG AAT TTT      2559
Glu Lys Glu Thr Met Asp Asn Ala Arg Ala Phe Cys Lys Arg Asn Phe
        800                 805                 870

GGT GAT CTT GTT TCT ATT CAA AGT GAA AGT GAA AAG AAG TTT CTA TGG      2607
Gly Asp Leu Val Ser Ile Gln Ser Glu Ser Glu Lys Lys Phe Leu Trp
    815                 820                 825

AAA TAT GTA AAC AGA AAT GAT GCA CAG TCT GCA TAT TTT ATT GGT TTA      2655
Lys Tyr Val Asn Arg Asn Asp Ala Gln Ser Ala Tyr Phe Ile Gly Leu
830                 835                 840                 845

TTG ATC AGC TTG GAT AAA AAG TTT GCT TGG ATG GAT GGA AGC AAA GTG      2703
Leu Ile Ser Leu Asp Lys Lys Phe Ala Trp Met Asp Gly Ser Lys Val
                850                 855                 860

GAT TAC GTG TCT TGG GCC ACA GGT GAA CCC AAT TTT GCA AAT GAA GAT      2751
Asp Tyr Val Ser Trp Ala Thr Gly Glu Pro Asn Phe Ala Asn Glu Asp
            865                 870                 875

GAA AAC TGT GTG ACC ATG TAT TCA AAT TCA GGG TTT TGG AAT GAC ATT      2799
Glu Asn Cys Val Thr Met Tyr Ser Asn Ser Gly Phe Trp Asn Asp Ile
        880                 885                 890

AAC TGT GGC TAT CCA AAC GCC TTC ATT TGC CAG CGA CAT AAC AGT AGT      2847
Asn Cys Gly Tyr Pro Asn Ala Phe Ile Cys Gln Arg His Asn Ser Ser
    895                 900                 905

ATC AAT GCT ACC ACA GTT ATG CCT ACC ATG CCC TCG GTC CCA TCA GGG      2895
Ile Asn Ala Thr Thr Val Met Pro Thr Met Pro Ser Val Pro Ser Gly
910                 915                 920                 925

TGC AAG GAA GGT TGG AAT TTC TAC AGC AAC AAG TGT TTC AAA ATC TTT      2943
Cys Lys Glu Gly Trp Asn Phe Tyr Ser Asn Lys Cys Phe Lys Ile Phe
                930                 935                 940

GGA TTT ATG GAA GAA GAA AGA AAA AAT TGG CAA GAG GCA CGA AAA GCT      2991
Gly Phe Met Glu Glu Glu Arg Lys Asn Trp Gln Glu Ala Arg Lys Ala
            945                 950                 955

TGT ATA GGC TTT GGA GGG AAT CTG GTC TCC ATA CAA AAT GAA AAA GAG      3039
Cys Ile Gly Phe Gly Gly Asn Leu Val Ser Ile Gln Asn Glu Lys Glu
        960                 965                 970

CAA GCA TTT CTT ACC TAT CAC ATG AAG GAC TCC ACT TTC AGT GCC TGG      3087
Gln Ala Phe Leu Thr Tyr His Met Lys Asp Ser Thr Phe Ser Ala Trp
    975                 980                 985
```

```
ACT GGG CTG AAT GAT GTC AAT TCA GAA CAC ACG TTC CTT TGG ACG GAT      3135
Thr Gly Leu Asn Asp Val Asn Ser Glu His Thr Phe Leu Trp Thr Asp
990                 995                 1000                1005

GGA CGA GGA GTC CAT TAC ACA AAC TGG GGG AAA GGT TAC CCT GGT GGA      3183
Gly Arg Gly Val His Tyr Thr Asn Trp Gly Lys Gly Tyr Pro Gly Gly
                1010                1015                1020

AGA AGA AGC AGT CTT TCT TAT GAA GAT GCT GAC TGT GTT GTT ATT ATT      3231
Arg Arg Ser Ser Leu Ser Tyr Glu Asp Ala Asp Cys Val Val Ile Ile
            1025                1030                1035

GGA GGT GCA TCA AAT GAA GCA GGA AAA TGG ATG GAT GAT ACC TGC GAC      3279
Gly Gly Ala Ser Asn Glu Ala Gly Lys Trp Met Asp Asp Thr Cys Asp
        1040                1045                1050

AGT AAA CGA GGC TAC ATA TGC CAG ACA CGA TCC GAC CCT TCC TTG ACT      3327
Ser Lys Arg Gly Tyr Ile Cys Gln Thr Arg Ser Asp Pro Ser Leu Thr
    1055                1060                1065

AAT CCT CCA GCA ACG ATT CAA ACA GAT GGC TTT GTT AAA TAT GGC AAA      3375
Asn Pro Pro Ala Thr Ile Gln Thr Asp Gly Phe Val Lys Tyr Gly Lys
1070                1075                1080                1090

AGC AGC TAT TCA CTC ATG AGA CAA AAA TTT CAA TGG CAT GAA GCG GAG      3423
Ser Ser Tyr Ser Leu Met Arg Gln Lys Phe Gln Trp His Glu Ala Glu
                1095                1100                1105

ACA TAC TGC AAG CTT CAC AAT TCC CTT ATA GCC AGC ATT CTG GAT CCC      3471
Thr Tyr Cys Lys Leu His Asn Ser Leu Ile Ala Ser Ile Leu Asp Pro
            1110                1115                1120

TAC AGT AAT GCA TTT GCG TGG CTG CAG ATG GAA ACA TCT AAT GAA CGT      3519
Tyr Ser Asn Ala Phe Ala Trp Leu Gln Met Glu Thr Ser Asn Glu Arg
        1125                1130                1135

GTG TGG ATC GCC CTG AAC AGT AAC TTG ACT GAT AAT CAA TAC ACT TGG      3567
Val Trp Ile Ala Leu Asn Ser Asn Leu Thr Asp Asn Gln Tyr Thr Trp
    1140                1145                1150

ACT GAT AAG TGG AGG GTG AGG TAC ACT AAC TGG GCT GCT GAT GAG CCC      3615
Thr Asp Lys Trp Arg Val Arg Tyr Thr Asn Trp Ala Ala Asp Glu Pro
1155                1160                1165                1170

AAA TTG AAA TCA GCA TGT GTT TAT CTG GAT CTT GAT GGC TAC TGG AAG      3663
Lys Leu Lys Ser Ala Cys Val Tyr Leu Asp Leu Asp Gly Tyr Trp Lys
                1175                1180                1185

ACA GCA CAT TGC AAT GAA AGT TTT TAC TTT CTC TGT AAA AGA TCA GAT      3711
Thr Ala His Cys Asn Glu Ser Phe Tyr Phe Leu Cys Lys Arg Ser Asp
            1190                1195                1200

GAA ATC CCT GCT ACT GAA CCC CCA CAA CTG CCT GGC AGA TGC CCG GAG      3759
Glu Ile Pro Ala Thr Glu Pro Pro Gln Leu Pro Gly Arg Cys Pro Glu
        1205                1210                1215

TCA GAT CAC ACA GCA TGG ATT CCT TTC CAT GGT CAC TGT TAC TAT ATT      3807
Ser Asp His Thr Ala Trp Ile Pro Phe His Gly His Cys Tyr Tyr Ile
    1220                1225                1230

GAG TCC TCA TAT ACA AGA AAC TGG GGC CAA GCT TCT CTG GAA TGT CTT      3855
Glu Ser Ser Tyr Thr Arg Asn Trp Gly Gln Ala Ser Leu Glu Cys Leu
1235                1240                1245                1250
```

```
CGA ATG GGT TCC TCT CTG GTT TCC ATT GAA AGT GCT GCA GAA TCC AGT      3903
Arg Met Gly Ser Ser Leu Val Ser Ile Glu Ser Ala Ala Glu Ser Ser
                1255                1260                1265

TTT CTG TCA TAT CGG GTT GAG CCA CTT AAA AGT AAA ACC AAT TTT TGG      3951
Phe Leu Ser Tyr Arg Val Glu Pro Leu Lys Ser Lys Thr Asn Phe Trp
            1270                1275                1280

ATA GGA TTG TTC AGA AAT GTT GAA GGG ACG TGG CTG TGG ATA AAT AAC      3999
Ile Gly Leu Phe Arg Asn Val Glu Gly Thr Trp Leu Trp Ile Asn Asn
        1285                1290                1295

AGT CCG GTC TCC TTT GTC AAC TGG AAC ACA GGA GAT CCC TCT GGT GAA      4047
Ser Pro Val Ser Phe Val Asn Trp Asn Thr Gly Asp Pro Ser Gly Glu
    1300                1305                1310

CGG AAT GAT TGT GTG ACT TTA CAT GCG TCT TCT GGG TTT TGG AGT AAT      4095
Arg Asn Asp Cys Val Thr Leu His Ala Ser Ser Gly Phe Trp Ser Asn
1315                1320                1325                1330

ATT CAC TGT TCT TCC TAC AAA GGA TAT ATT TGT AAA AGA CCA AAA ATT      4143
Ile His Cys Ser Ser Tyr Lys Gly Tyr Ile Cys Lys Arg Pro Lys Ile
                1335                1340                1345

ATT GAT GCT AAA CCT ACT CAT GAA TTA CTT ACA ACA AAA GCT GAC ACA      4191
Ile Asp Ala Lys Pro Thr His Glu Leu Leu Thr Thr Lys Ala Asp Thr
            1350                1355                1360

AGG AAG ATG GAC CCT TCT AAA CCG TCT TCC AAC GTG GCC GGA GTA GTC      4239
Arg Lys Met Asp Pro Ser Lys Pro Ser Ser Asn Val Ala Gly Val Val
        1365                1370                1375

ATC ATT GTG ATC CTC CTG ATT TTA ACG GGT GCT GGC CTT GCC GCC TAT      4287
Ile Ile Val Ile Leu Leu Ile Leu Thr Gly Ala Gly Leu Ala Ala Tyr
    1380                1385                1390

TTC TTT TAT AAG AAA AGA CGT GTG CAC CTA CCT CAA GAG GGC GCC TTT      4335
Phe Phe Tyr Lys Lys Arg Arg Val His Leu Pro Gln Glu Gly Ala Phe
1395                1400                1405                1410

GAA AAC ACT CTG TAT TTT AAC AGT CAG TCA AGC CCA GGA ACT AGT GAT      4383
Glu Asn Thr Leu Tyr Phe Asn Ser Gln Ser Ser Pro Gly Thr Ser Asp
                1415                1420                1425

ATG AAA GAT CTC GTG GGC AAT ATT GAA CAG AAT GAA CAC TCG GTC ATC      4431
Met Lys Asp Leu Val Gly Asn Ile Glu Gln Asn Glu His Ser Val Ile
            1430                1435                1440

TAG TACCTCAATG CGATTCTGAG ATATTTGAAT TTCATAAAAT TGTAACTGAA          4484
End

ATTTAAAATT TTTAGTTCAA TGTGATTGTT TTCTTTAAAA TGAGTACTGA ATTGTACTGG   4544

TCTGTCCTTT TTTCCTTTGC CTAATTGAAG AAATAATTGC TTGTTTTCTA GCCTGGCAAG   4604

ATATTTTCAT AAAAGAGGGA TAACAATGCT GATTACTACC TTTTAAAATA TTTTAGATAA   4664

ATGCACAGCA CCACAGCACC ACATCTAAGC ATTAGTGATG GGTAGCTGAT GTCAGCTTCA   4724

TGTGGATTTT AAGCACTCTA GAAACAATGA AGCTTCTTGG CATATTTTAA GGAGCTCCCA   4784

AAATGTGTTA CCTATTAAAT TGTAACTCAG CAAGTAGAAG ACCATTTGAA AAGTCAGGTA   4844

CAAATTTCCT CAAGTGGCAT AAAAATGTAG TCAGTTTTCT CTTTTACCAG TTTTTATTTC   4904
```

CACTCCAATT ATTTAGAACT TTATTTGTAC ATGTGCAGAA GAATAAGGCA GCTGAGAATC 4964

TTGTTTCCCC CAAGAGAGTT TTACAGGCTG AGTGTTGCAA ATGTGTTCTT TGTCCTGTTA 5024

TATGTATATC AGGAATACAA GGATGTGAAA TAAAACTGTA AATTTGCATA ACTGGATGTA 5084

CTTAGATAAT GTGAAATAAA CATTAAAGAC AAGGTCTATT TTTAATAAAA AAAAAAAAAA 5144

A 5145

What is claimed is:

1. A soluble recombinant peptide comprising at least one carbohydrate recognition domain, derived from an extracellular portion of mannose receptor protein, said peptide lacking the mannose receptor protein transmembrane and cytoplasmic regions, said peptide being capable of specifically targeting cells expressing mannose, N-acetylglucosamine, or fucose.

2. The peptide of claim 1 comprising a sequence with greater than 75% homology to a fragment of at least 150 contiguous amino acids of mannose receptor protein.

3. The peptide of claim 1 comprising at least 150 contiguous amino acids of mannose receptor protein.

4. The peptide of claim 1 comprising a sequence with greater than 75% homology to a fragment of at least 300 contiguous amino acids of mannose receptor protein.

5. The peptide of claim 1 comprising at least 300 contiguous amino acids of mannose receptor protein.

6. The peptide of claim 1 comprising at least one complete carbohydrate recognition domain from mannose receptor protein, shown in Fig. 3.

7. The peptide of claim 1 comprising at least two complete carbohydrate recognition domains from mannose receptor protein, shown in Fig. 3.

8. The peptide of claim 1 comprising at least two copies of a carbohydrate recognition domain from mannose reception protein, shown in Fig. 3.

9. The peptide of claim 1 further comprising a segment capable of fixing complement.

10. The peptide of claim 9 wherein said peptide comprises a complement-fixing region of immunoglobulin heavy chain.

11. The peptide of claim 1 further comprising a cytotoxin.

12. The peptide of claim 11 wherein said cytotoxin is AZT, ricin, pertussis, or cholera toxin.

13. Engineered nucleic acid encoding the peptide of claim 1.

14. The engineered nucleic acid of claim 13 wherein said nucleic acid is cDNA.

15. A nucleic acid fragment substantially corresponding to at least 450 contiguous bases of the nucleic acid encoding the soluble extracellular fragment of mannose receptor protein (SEQ ID NO: 1), deposited in the ATCC as ATCC No. 68430, said fragment encoding a soluble peptide capable of specifically targeting cells expressing mannose, N-acetylglucosamine, or fucose.

16. Engineered nucleic acid substantially corresponding to the nucleic acid encoding the soluble

extracellular fragment of mannose receptor protein (SEQ ID NO: 1), deposited in the ATCC as ATCC No. 68430.

17. An expression vector comprising the engineered nucleic acid of claim 13.

18. A recombinant cell comprising the engineered nucleic acid of claim 13.

19. A therapeutic agent comprising a therapeutically effective amount of the peptide of claim 1 administered in a pharmaceutically acceptable carrier substance.

20. The therapeutic agent of claim 19 comprising a lyposome coated with the soluble peptide.

21. A method for treating an animal infected with a bacterium, a fungus, or a virus, said method comprising the steps of

providing the therapeutic agent of claim 19, and

administering to the animal a therapeutically effective amount of said agent.

22. The method of claim 21 wherein said animal is human.

23. The method of claim 22 wherein said peptide is administered by application of a powder or a lotion comprising said peptide to the foot.

24. The method of claim 21 wherein said agent lowers the level of infection of eucaryotic cells by said bacterium, fungus, or virus.

25. A purified antibody specifically recognizing the peptide of claim 1.

26. A method for diagnosing, in an animal, an infection by a bacterium, a fungus, or a virus, said method comprising the steps of

providing a sample of serum from the animal,

contacting said serum with an effective amount of the peptide of claim 1, and

measuring the amount of binding of said peptide to said serum sample.

27. The method of claim 26 wherein said animal is human and said infection is by a fungus.

28. The method of claim 26 wherein said measuring step comprises determining the amount of binding of an antibody specific for said peptide to a serum sample containing said peptide.

29. The method of claim 26 wherein said measuring step comprises fluorescence detection.

30. A purified soluble peptide comprising the extracellular portion of mannose receptor protein, said peptide lacking the mannose receptor protein transmembrane and cytoplasmic regions.

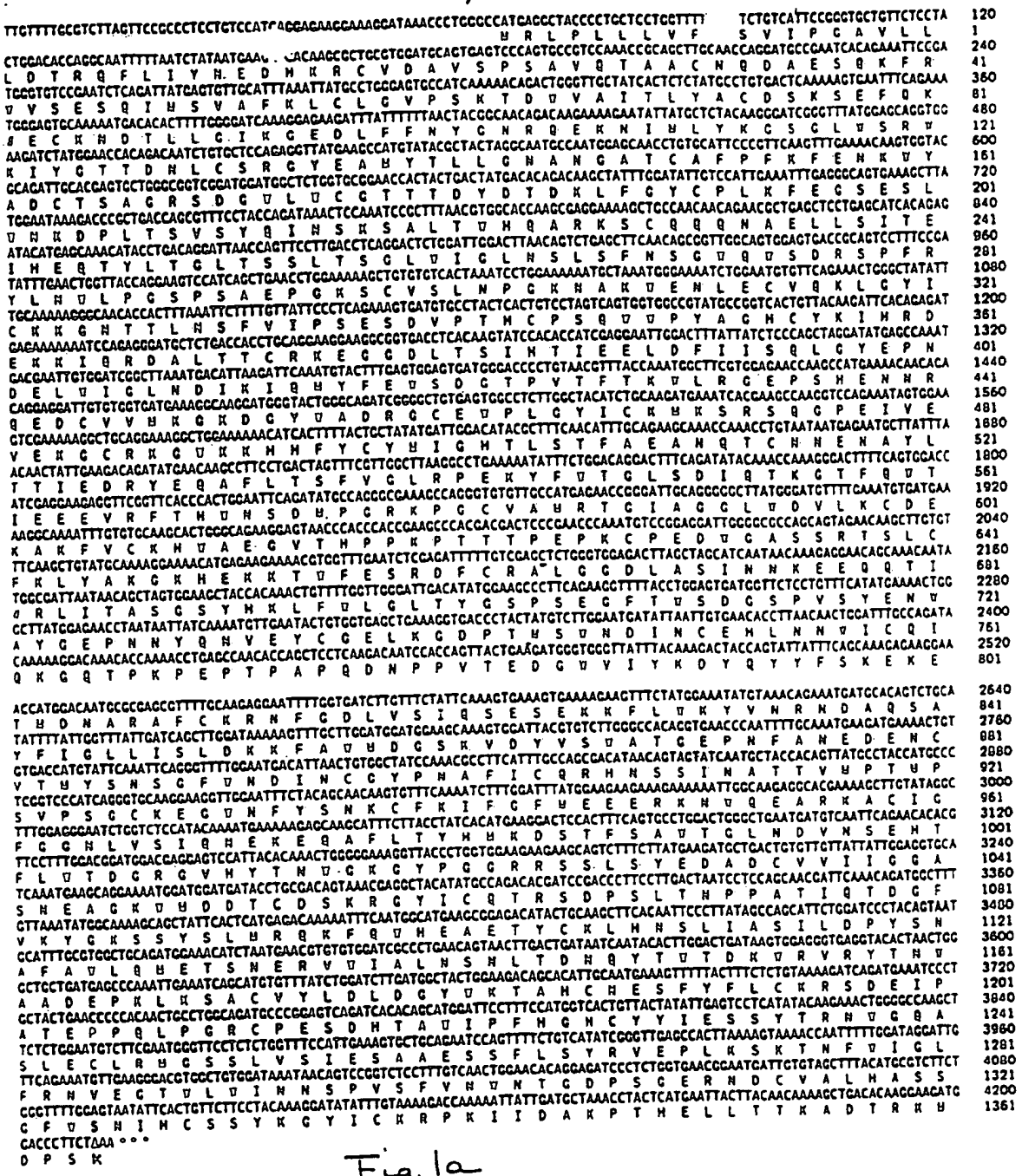

Fig. 1a

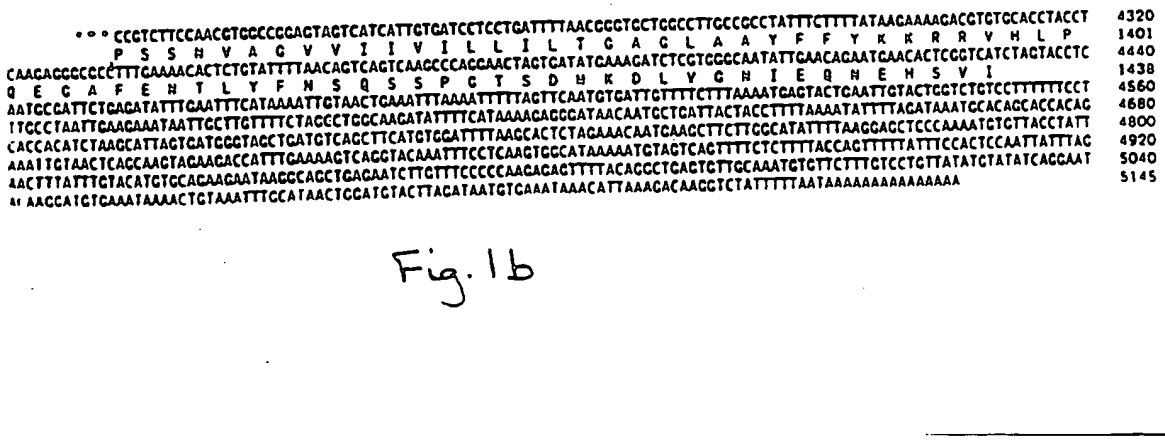

Fig. 1b

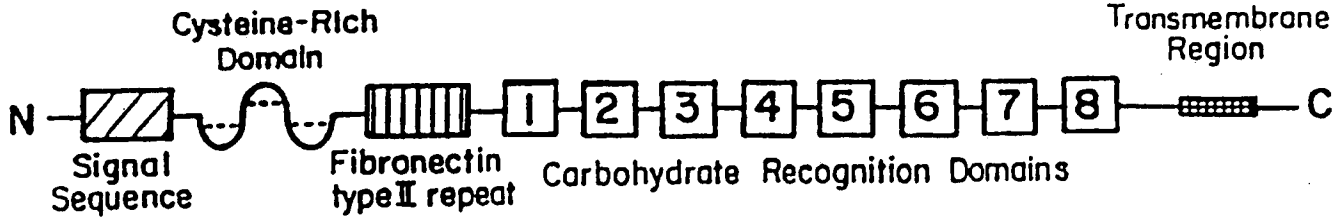


Fig. 2

```
CRD 1  (189)                    GYCPLKF..EGSESLWNKDPLTSVS....YQINSKSALTWHQARKSCQQQNAELLSITEIHEQTYL.TG.LT.S.SLTS.GLWIGLNSLSFNSGW.QWSDRSPFRYLNWLPGSPSAEPG.....K.SCVSLNPG...KNAKWENLECVQK
CRD 2  (324)    KGNTTLN.SF.....VIPSE.SDVPTHC....PSQWWPYACHCYKI....HRDEKKIQRD..ALTTCRKECGQLTSIHTIEELDFI.ISQLGY..EP.NDELWIGLNDIKIQWY.FEWSDGTPVTFTKWLRGEPSHE....NKRQEDCVVMK.G...KDGYWADRGCEMP
CRD 3  (470)       WKSRSQGP...EIPEIVEVEK.GCRK....GWKKHHFYCYWI..GHTLSTFAE.....ANQTCNNENAYLTTIEDRYEQAFL.TSFVGL..RP.EKYFWTGLSDIQTKGT.FQWTIEEEVRFTHWNSDWP.......GRKPGCVAWRTG..IACGLWDVLKCDEK
CRD 4  (609) HWAEGVTHPP........KPTTTPEP.K..CPED..WGASSRTSLCFK...LYAKGKHEKKTWFESRDFCRALGGDLASIHNKEEQQTIWRLI.TASGS.YHKLFWLGLTYGSPSEG.FTWSDGSPVSYENWAYGEPN.....NYQNVEYCGELK.G..DPTWSWNDINCEHL
CRD 5  (761)  IQXGQTPKP........EPTPAPQDNPPV.TED...GWVIYKDY..:-...QYYFSKEKETWDN.ARAFCKRNFGDLVSIQSESEKKFLWKYVNRN..DAQSAYF.IGL..LISLDKK.FAWWDGSKVDYVSWATGEPN.....FANEDENCVTMYS....NSGFWNDINCGYP
CRD 6  (905)  RHNSSINAT.......TVWPTWPSVPS.GCKE....GWNFYSNKCFKIFGFWEE...ERKNWQEARKACIGFGGNLVSIQNEKEQAFL.TYHWKD..STFSA..WTGLNDVNSEHT.FLWTDGRGVHYTNWGKGYPGGRRSSLSYEDADCVVIIGGASNEAGKWNDDTCDSK
CRD 7 (1063)           TRSD.......PSLTWPPATIQT.D...GFVKYGKSS......YSLWRQ.KFQWHEAETYCKLHNSLIASILDPYSNAFAW.LQWET..SN.ERV.WIALNS.NLTDNQYTWIDKSKRVRYTNWAADEPKGX.SA.......CVYL.....DLDGYWKTAHCNES
CRD 8 (1196)            RSDEIPATEPPQL...PGR.CPESDHTAWIPFHGHC......YYIESSYTRNWGQASLECLRWGSSLVSIESAAESSFLS.YRVEPLKSKTN..FWIGLFR.NVEGTW.LWINNSPVSFVNWNTGDPS.......GERNDCVTLHA....SSGFWSNIHCSSY
```

Fig. 3

# INTERNATIONAL SEARCH REPORT

PCT/US91/08320

International Application No.

## I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): A61K 37/02; C07K 7/00, 13/00, 15/00

U.S. Cl. 530/300, 350, 395; 424/88; 514/8, 12

## II. FIELDS SEARCHED

Minimum Documentation Searched [7]

| Classification System | Classification Symbols |
|---|---|
| U.S. Cl. | 530/300, 350, 395; 424/88; 514/8, 12 |

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

APS, CAS, BIOSIS, PIR, SWISS-PROT; Search terms: Mannose receptor protein; carbohydrate recognition domains (CRDs)

## III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

| Category [10] | Citation of Document, [11] with indication, where appropriate, of the relevant passages [12] | Relevant to Claim No. [13] |
|---|---|---|
| Y | THE JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 262, No. 20, issued 15 JULY 1987, HALBERG ET AL., "Major and Minor Forms Of The Rat Liver Asialoglycoprotein Receptor Are Independent Galactose-Binding Proteins", pages 9828-9838. See the entire document. | 1-12, 30 |
| Y | JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 265, No. 21, issued 25 JULY 1990, TAYLOR ET AL. "Primary Structure Of The Mannose Receptor Contains Multiple Motifs Resembling Carbohydrate Recognition Domains", pages 12156-12162. See the entire document. | 1-12, 30 |
| Y | THE JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 262, No. 21, issued 25 JULY 1987, LENNARTZ ET AL., "Isolation And Characterization Of A Mannose-Specific Endocytosis Receptor From Human Placenta, "Pages 9942-9944, See especially the abstract, Figure 2, Table 1, page 9944, column 1. | 1-12, 30 |

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

## IV. CERTIFICATION

| Date of the Actual Completion of the International Search | Date of Mailing of this International Search Report |
|---|---|
| 29 January 1992 | 27 FEB 1992 |
| **International Searching Authority** | **Signature of Authorized Officer** |
| ISA/US | Lisa T. Bennett |

Form PCT/ISA/210 (second sheet) (Rev. 11-87)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

| Category * | Citation of Document, with indication, where appropriate, of the relevant passages | Relevant to Claim No |
|---|---|---|
| Y | JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 261, No. 16, issued 05 JUNE 1986, HALTIWANGER ET AL., "The Isolation Of A Rat Alveolar Macrophage Lectin", Pages 7440-7444, See especially the Abstract and Table II. | 1-12, 30 |
| Y | JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 263, No. 20, issued 15 JULY 1988, DRICKAMER, Two Distinct Classes of Carbohydrate Recognition Domains In Animal Lectins", Pages 9557-9560, See especially page 9557, column 1 and 9558, column 1, paragraph 3-column 2. | 1-12, 19-20 30, 25 |
| Y | BIOCHEMICAL JOURNAL, Volume 245, Issued in 1987, LENNARTZ ET AL., "Isolation and Characterization Of A Mannose-Specific Endocytosis Receptor From Rabbit Alveolar Macrophages", Pages 705-711, See especially the Abstract, page 705, column 1, figure 2, page 708. | 1-12, 30 19-20, 25 |
| Y | BIOCHEMICAL JOURNAL, Volume 262, Issued 1989, CHILDS ET AL., "Neoglycolipids As Probes Of Oligosaccharide Recognition By Recombinant And Natural Mannose-Binding Proteins Of Rat and Man", Pages 131-138, See especially Abstract; pages 131, Column 2,-page 131, column 1 and page 136. | 1-12, 30 19-20, 25 |
| Y | JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 259, No. 3, issued 10 FEBRUARY 1984, LARGENT ET AL., "Carbohydrate-Specific Adhesion Of Alveolar Macrophages To Mannose-Derived Surfaces", Pages 1764-1769, See especially the Abstract and page 1764, column 2, figures 1-3. | 1-12, 30 19-20, 25 |